CSC321 Programming Assignment 2

Vitaly Topekha

February 28, 2018

1 A. Colourization as Regression

1. There are 6 convolution layers:

	Fileter Size / Kernel	# of Filters
downconv1	3x3	32
downconv2	3x3	64
rfconv	3x3	64
upconv1	3x3	32
upconv2	3x3	3
finalconv	3x3	3

2. The resulting colours are not similar to the ground truth images. The background colours are not precise and are wrong. The supposedly blue colour in each image becomes either greenish or brownish and the brightness of the green colours is off.

3. RGB values of the same colours can be diffrerent depending on illumination. For instance, two cars of the same colour can have different rgb values in different lightning.

4. Given very confident or very wrong predictions, regression will penalize them. This leads the predicted colours to be not vibrant, and similar across all images.

2 B. Colourization as Classification

2. The resulting colours are much brighter and more accurate then the previous model. However the images appear to be more blurry and less precise.

3 C. Skip Connections

3. Way better than previous model. Sharpness has improved. This model is able to pick up more little details from the images, unlike previous models.

   CNN - Val Loss: 1.5881, Val Acc: 41.1%
   UNet - Val Loss: 1.3659, Val Acc: 48.0%

   The skip connection improves the loss and accuracy.
   Although some information is lost during pooling, the skip connection helps to preserve the lost information while in reconstruction during the up-sampling.

4 D. Dilated Convolution

1. (a) 3x3 convolution: 9 weights and receptive field of size 9 pixels
   (b) 5x5 convolution: 25 weights and receptive fields of size 25 pixels
   (c) 3x3 with dilation 1 convolution: 9 weights and receptive fields of size 25 pixels

2. Using dilation in first few layers could lead to loss of some features while learning localized features like edges. Using dilation in the middle layers, we can get some invariance to the scale of the image or learn features of the image in a bigger picture with less weights. It also gives invariance to shifting in 2D.

5 E. Visualizing Intermediate Activations

1. The activations of the last few layers are more pixelated, its not clear what's being highlighed in activations. For the activations of first few layers, they show the outline of the image under different gradients or lightings, the outline of the horses is way more vibrant.

2. In the last layer of the UNet, the activations show more detailed resemblance to the outline of the original image in comparison to the CNN activations in the last layer. The activations of the last layer of UNet can show details such as a person on the horse, while the activations of the last layer of CNN shows the outline of an object which is hardly recognizable.

6 F. Conceptional Problems

1. (a) and (b) could be used for new training data. (c) and (d) does not make any difference due to the of max pooling. (e) does not help because we are learning colour of horses.